

Abstract of the Disclosure

A method identifies a talking head or principal cast in a compressed video by first segmenting the video into shots. Motion activity descriptors are extracted from each of the shots, and combined into a shot motion activity descriptor. A distance between the shot motion activity descriptor and a template motion activity descriptor is measured. The template motion activity descriptor is obtained from a training video. If the measured distance is less than a predetermined threshold, then the shot is identified as including a talking head.